



Datová úložiště v MetaCentru a okolí II

David Antoš
Oddělení datových úložišť

Přehled

- ▶ úložiště v MetaCentru
 - ▶ proč (většinou) nemám počítat nad daty ze /storage?
 - ▶ proč je v Brně plné pole?
 - ▶ jak mám přesouvat větší data?
- ▶ úložiště v okolí
 - ▶ hierarchická datová úložiště CESNET a CERIT-SC
 - ▶ na co je mám použít v MetaCentru?
 - ▶ na co je mohu používat i jinde?

Úložiště v MetaCentru

- ▶ scratche
- ▶ svazky /storage
 - ▶ těch máme hodně
- ▶ domovské adresáře na clusterech
- ▶ /software, /packages (AFS)
- ▶ hierarchické úložiště (HSM) pro archivní data

Úložiště v MetaCentru

- ▶ scratche
- ▶ svazky /storage
 - ▶ těch máme hodně
- ▶ domovské adresáře na clusterech
- ▶ /software, /packages (AFS)
- ▶ hierarchické úložiště (HSM) pro archivní data

těch ale je!

- ▶ mají různé vlastnosti, které potřebujeme
- ▶ důsledek: je potřeba o nich vědět

Scratch

- ▶ (lokální) úložiště dat při běhu úlohy na výpočetním uzlu
- ▶ použití: stage-in/stage-out
- ▶ typy
 - ▶ `/scratch.ssd` (SSD, malé, velmi rychlé)
 - ▶ `/scratch` (lokální rotační HD)
 - ▶ `/scratch.shared` (sdílený na clusteru)
 - ▶ volbou typu `ssd` nebo `shared` omezíte možné stroje
- ▶ novinka: do `/scratch*/$USER` nelze zapisovat (práva 555)
 - ▶ úlohy mohou používat `/scratch*/$USER/job_JOBID`
 - ▶ jeho hodnotu zjistí v proměnné `$SCRATCHDIR`
- ▶ vychovaná úloha po sobě `scratch` smaže
 - ▶ po nevychovaných úlohách se `scratch` maže po 14 dnech

Scratch 2

- ▶ při zadání úlohy lze vyžádat kapacitu a případně typ
- ▶ např. `-l scratch=10gb:ssd`
- ▶ `-l scratch=10gb` vybere nejrychlejší dostupný typ
- ▶ v plánovači jsou skripty pro přesuny dat do scratche a zpět
 - ▶ viz dokumentace na wiki MetaCentra

Svazky /storage

- ▶ disková pole „v každém městě“
- ▶ adresáře uživatelů /storage/**město**/home/**uživatel**
- ▶ jsou dostupné ze všech strojů s vlastností nfs4
 - ▶ tj. téměř odevšad
- ▶ autentizace Kerberem – „potřebujete lístky“
- ▶ dají se připojit i na vlastní pracovní stanici přes NFSv4
- ▶ na jednotlivá úložiště můžete přistupovat i přes ssh, scp, SFTP
 - ▶ v návodech je přehled, jak se jmenují stroje

Svazky /storage

Your Home Directories

```
-----
```

Size	Avail	S.Quota	Used	Directory
85T	5.9T	5T	3G	/storage/brno1/home/antos
110T	4.7T	3T	278M	/storage/brno2/home/antos
260T	35T	1T	256k	/storage/brno3-cerit/home/anto..
44T	32T	3T	28k	/storage/budejovice1/home/anto..
374T	356T	1T	0	/storage/jihlava1-cerit/home/a..
88T	88T	3T	4k	/storage/ostrava1/home/antos
44T	31T	1T	1M	/storage/plzen1/home/antos
51T	33T	3T	84k	/storage/praha1/home/antos

/storage – doporučení

- ▶ který /storage mám používat
 - ▶ nějaký volný
 - ▶ pro velké soubory to není kritické
 - ▶ pro malé soubory fyzicky nejbližší ke clusteru
- ▶ fyzicky nejbližší /storage je použit jako /home clusteru
- ▶ pro co se /storage nehodí
 - ▶ pro přímé výpočty nad daty
 - ▶ s výjimkou sekvenčního čtení velkých souborů
 - ▶ archivy, dlouho nepoužívaná data
 - ▶ na to máme hierarchické úložiště

/software, /packages (/afs)

- ▶ dnes již výhradně pro instalace software
- ▶ autentizace Kerberem
- ▶ umožňuje velmi detailní správu přístupových práv
- ▶ lze připojit k vlastní stanici

Přístupová práva, kvóty

- ▶ přístupová práva
 - ▶ klasická unixová (standardně rwxr-xr-x)
 - ▶ ACL na AFS
 - ▶ ACL na NFS
- ▶ kvóty
 - ▶ soft quota vs. limit
 - ▶ přehled je na osobní stránce
 - ▶ na webu je autoritativní zdroj
 - ▶ na starších systémech program quota dává nesmyslné výstupy
 - ▶ lze požádat o změnu

Přesuny dat, doporučení

- ▶ „velká data“ – stovky GB a více
- ▶ pokud potřebujete větší přesuny dat, vezměte si interaktivní úlohy (ne přímo na frontendu)
 - ▶ lze použít cp mezi připojenými adresáři
 - ▶ lze kopírovat data přímo z úložiště
- ▶ pro přesuny opravdu velkých dat (více TB) se domluvte s uživatelskou podporou
- ▶ archivní data přesuňte na hierarchická úložiště

Hierarchická úložiště

- ▶ celková hrubá kapacita cca 21 PB
- ▶ Plzeň (od pol. 2012)
 - ▶ v areálu ZČU
 - ▶ cca 500 TB disků, 4800 TB pásek, 3584 TB SGI COPAN
 - ▶ SGI Data Migration Facility, CXFS
- ▶ Jihlava (2013)
 - ▶ Krajský úřad Jihlava
 - ▶ 1041 TB disků, 3744 TB pásek, 2934 TB MAID
- ▶ Brno (2013)
 - ▶ rektorát VUT Brno
 - ▶ 498 TB disků, 3500 TB pásek, 2116 TB MAID
 - ▶ řešení IBM založené na GPFS a TSM

Úložiště jsou hierarchická

- ▶ vrstvy médií různé kapacity a rychlosti
 - ▶ rychlé disky/pomalejší disky/MAID/pásky
 - ▶ drahý provoz → levnější provoz
- ▶ a automatizovaný systém pro přesuny dat mezi nimi
 - ▶ déle nepoužívaná data do pomalejších vrstev
 - ▶ pro uživatele transparentní, resp. téměř transparentní
 - ▶ přístup k dlouho nepoužitému souboru trvá déle
 - ▶ uživatelům to obvykle nevadí

Možnosti využití datových úložišť

- ▶ zálohy
 - ▶ uživatelé mají primární data u sebe
 - ▶ na úložiště odkládají zálohu pro případ havárie
- ▶ archivace
 - ▶ uživatelé na úložiště odkládají cenná primární data
 - ▶ data nejsou často využívána
 - ▶ uživatelé nemají prostředky pro jejich uchování
- ▶ sdílení dat
 - ▶ distribuovaný tým potřebuje společně pracovat nad většími objemy dat, případně je zveřejňovat
- ▶ „něco jiného“
 - ▶ distribuce obsahu, jiné speciální aplikace

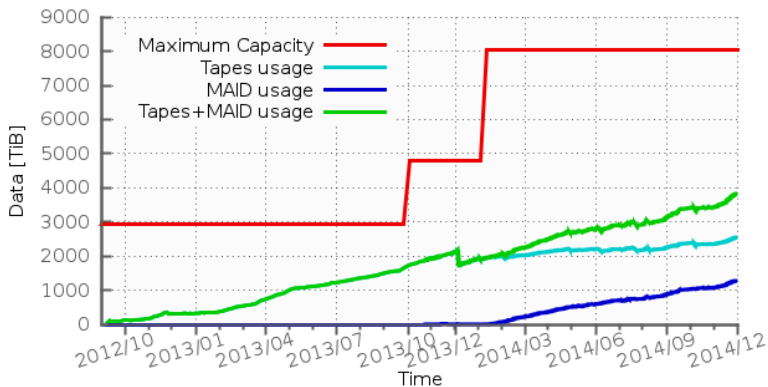
Knihovna v Plzni



Zaplnění úložiště v Plzni

Data stored on tapes/MAID

CESNET Pilsen storage, 2014-11-30 06:16:02



Přenos dat na HSM

- ▶ souborové
 - ▶ NFSv4 – přímo připojeno do MetaCentra
 - ▶ rsync, scp, FTPS
 - ▶ Globus – vysokorychlostní přenosy řízené webovým rozhraním
- ▶ speciální
 - ▶ gridové úložiště v systému dCache
 - ▶ FileSender – jednorázový přenos souborů
 - ▶ <http://filesender.cesnet.cz>
 - ▶ ownCloud

FileSender

- ▶ webová služba pro jednorázový přenos (velkých) souborů
 - ▶ velkých: aktuálně 500 GB
- ▶ přispíváme i k jeho vývoji
- ▶ alespoň jedna strana komunikace musí být oprávněný uživatel infrastruktury
 - ▶ autentizace federací eduID.cz
- ▶ oprávněný uživatel může nahrát soubor a poslat mu oznámení
- ▶ lze poslat komukoli pozvánku
- ▶ připravujeme upload z dávkové úlohy

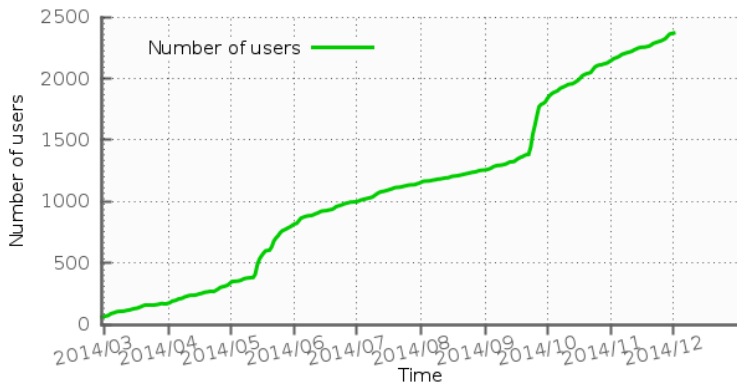
ownCloud

- ▶ cloudové úložiště se synchronizací souborů
 - ▶ pro osobní počítače (Windows, Linux, Mac OS X)
 - ▶ mobilní Android, Apple (klient pod 1 Euro)
 - ▶ a webové rozhraní
- ▶ data se synchronizují přes úložiště
 - ▶ na počítači jsou i lokálně, na mobilní při otevření
- ▶ data lze sdílet
 - ▶ konkrétní osobě nebo „kdo zná odkaz“
- ▶ kalendář, kontakty
- ▶ registrace federací na <http://owncloud.cesnet.cz>
- ▶ standardní limit 100 GB na uživatele

ownCloud – počet uživatelů

CESNET ownCloud users

2014-12-01 05:16:01



Správa uživatelských účtů

- ▶ systém Perun
- ▶ uživatel (osoba) má (ideálně) jedinou identitu v infrastruktuře
- ▶ potřebujeme ověření identity a příslušnosti k oprávněné instituci
 - ▶ proč to? zodpovídáme v projektu za to, komu služby poskytujeme (akademické komunitě), řídíme mnoho různých služeb
- ▶ na identitu uživatele jsou navázány jednotlivé mechanismy přístupu ke zdrojům
- ▶ řízeno členstvím ve virtuálních organizacích

Jak se stát uživatelem – „nedělat nic“

- ▶ ownCloud, FileSender – přihlašte se federací
- ▶ uživatelé MetaCentra už úložiště používají přímo
 - ▶ přímo připojeno na clustery MetaCentra (NFSv4)
 - ▶ uživatelé MetaCentra tam mají adresáře
 - ▶ podporu zajišťuje MetaCentrum
 - ▶ tj. meta@cesnet.cz
- ▶ i nepřímo – zálohy z MetaCentra
- ▶ připojeno jako /storage/plzen2-archive/home/...
 - ▶ ale pozor, neukládejte data do home, ale do prostoru VO MetaCentrum

Jak se stát uživatelem – speciální požadavky

- ▶ pokud chcete využívat souborové protokoly ve větší míře
- ▶ napište nám na du-support@cesnet.cz
- ▶ domluvíme se, co potřebujete
 - ▶ opravdu to potřebujeme pochopit
 - ▶ kolik dat máte, jak s nimi pracujete, ...
- ▶ připravíme konfiguraci úložiště
- ▶ založíme virtuální organizaci
- ▶ pozveme vás do ní a povýšíme na administrátora
- ▶ dál si uživatele spravujete sami, vytváříte skupiny, ...
- ▶ ukládáte data přes FTP/rsync/NFSv4/...
 - ▶ my se staráme, aby na se vytvářely jejich repliky, ...

Shrnutí

- ▶ pro efektivní práci s daty je třeba mít představu o možnostech jejich ukládání
- ▶ máme hierarchická úložiště
 - ▶ pro archivní data
 - ▶ přímo přístupná z MetaCentra
 - ▶ uživatelskou podporu zajišťuje MetaCentrum standardními kanály
 - ▶ i pro další použití
 - ▶ podrobnější informace viz <http://du.cesnet.cz> a du-support@cesnet.cz